

Multicanonical distribution and the origin of power laws

G. L. Vasconcelos* and D. S. P. Salazar†

*Departamento de Física, Universidade Federal de Pernambuco, Brazil. and
Unidade de Educação a Distância e Tecnologia,
Universidade Federal Rural de Pernambuco, Brazil.*

Abstract

A multicanonical formalism is applied to the problem of statistical equilibrium in a complex system with a hierarchy of dynamical structures. At the small scales the system is in quasi-equilibrium and follows a Maxwell-Boltzmann distribution with a slowly fluctuating temperature. The probability distribution for the temperature is determined using Bayesian analysis and it is then used to average the Maxwell-Boltzmann distribution. The resulting energy distribution law is written in terms of generalized hypergeometric functions, which display power-law tails.

PACS numbers: Complex systems 89.75.-k, Classical ensemble theory 05.20.Gg

* giovani@df.ufpe.br

† dsps@df.ufpe.br

I. INTRODUCTION

Power law distributions occur in widely diverse physical systems spanning an impressive range of length scales [1]. The appearance of heavy-tailed distributions is often traced to the presence of hierarchical structures in the system [2], whose complex “interaction” may result in violation of the statistical independence of subsystems, thus leading to non-Gibbsian distributions [3]. Despite these insights and the many contributions to the problem [4], it is fair to say that the physical mechanisms behind the emergence of power-law distributions are not yet well understood.

Recently, we introduced [6] a general formalism to describe statistical equilibrium of complex systems with multiple scales where the probability distribution of states displays power-law tails. In this formalism, the hierarchical structure embedding the system of interest is effectively modeled as a set of nested “internal heat reservoirs,” where each “reservoir” is described by only one effective degree of freedom, namely, its “temperature.” On the basis of a few physically reasonable assumptions, it was possible to show that for a large class of systems the equilibrium distribution can be written explicitly in terms of certain generalized hypergeometric functions, which exhibit power law tails. This family of generalized hypergeometric (GHG) distributions includes, as its first two members, the Boltzmann-Gibbs distribution and the Tsallis distribution [7]. Higher-order members of the GHG family of distributions have been shown to describe remarkably well the statistics of velocity fluctuations in turbulence [5].

The GHG distribution represents a generalization of the canonical distribution for multiscale systems and hence it is also called multicanonical. The main purpose of the present paper is to give an alternative derivation, based on Bayesian analysis, of the multicanonical distribution. Because the derivation of the multicanonical distribution given here tries to parallel (whenever possible) the usual treatment of the canonical distribution, we shall begin our presentation by briefly reviewing the derivation of the canonical distribution.

II. THE CANONICAL DISTRIBUTION

Consider a system in thermal equilibrium at some temperature T_0 . We recall that in the canonical formalism the constant temperature constraint is enforced by embedding the

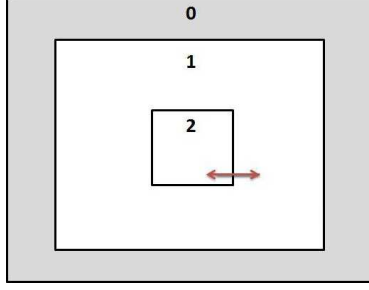


FIG. 1. A canonical system in thermal equilibrium at temperature T_0 . Arrows indicate energy exchange between the subsystems.

system is a much larger system (i.e., a heat bath) capable of giving it energy [8]. We designate the system of interest by the label 2 and the larger system embedding it by the label 1, with the combined system formed by subsystems 1 and 2 being given the label 0; see Fig. 1.

The energies of systems 1 and 2 will be denoted by E_k^1 and E_i^2 , respectively, where k and i represent the labels designating the possible states in each system. Let us denote by $p_0(E_{k,i}) = p_0(E_k^1 + E_i^2)$ the probability of finding system 0 in a state corresponding to energy $E_{k,i} = E_k^1 + E_i^2$. In view of the independence between systems 1 and 2, one can then write

$$p_0(E_k + E_i) = p_1(E_k)p_2(E_i), \quad (1)$$

where $p_1(E_k)$ and $p_2(E_i)$ are the probabilities of finding systems 1 and 2 in states with energy E_k and E_i , respectively. Taking the logarithm derivative of (1) with respect to E_k yields

$$\frac{\partial \ln p_0(E_i + E_k)}{\partial E_k} = \frac{d \ln p_1(E_k)}{d E_k} \equiv -\beta_1, \quad (2)$$

where β_1 may be a function of E_k but *not* of E_i . On the other hand, it is clear that

$$\frac{\partial \ln p_0}{\partial E_k} = \frac{\partial \ln p_0}{\partial E_i} = \frac{d \ln p_2(E_i)}{d E_i}. \quad (3)$$

Comparing (2) and (3), one then concludes that

$$\frac{d \ln p_2(E_i)}{d E_i} = -\beta_1. \quad (4)$$

Since β_1 does not depend on E_i , the preceding equation can be readily integrated, yielding

$$p_2(E_i|\beta_1) = \text{constant} \cdot \exp(-\beta_1 E_i). \quad (5)$$

Since the partition of system 0 into subsystems 1 and 2 is entirely arbitrary, the quantity β_1 must be the same for any partition one chooses. In other words, β_1 is a characteristic of system 0 only, that is,

$$\beta_1 = \beta_0, \quad (6)$$

which implies that

$$p_2(E_i|\beta_0) = A \exp(-\beta_0 E_i), \quad (7)$$

where A is a constant.

It should be evident that the discussion above is completely symmetrical with respect to labels 1 and 2, so that system 2 (irrespective of its size) obeys the same distribution law as system 1. It then follows from (1) and (7) that the probability of finding any subsystem of system 0 in a state with energy E is given by the Boltzmann-Gibbs (BG) distribution:

$$p(E|\beta_0) = \frac{g(E) \exp(-\beta_0 E)}{Z_0(\beta_0)}, \quad (8)$$

where $g(E)$ is the density of states and

$$Z_0(\beta) = \int_0^\infty g(E) \exp(-\beta E) dE \quad (9)$$

is the partition function

It is important to emphasize here that the key step in deriving (8) was the ability to partition the system into two independent subsystems of arbitrary sizes, so that each subsystem is described by the same distribution law. There are however many physical systems, where the relevant probability distributions depend on the scale at which the measurements are made. In such complex systems, the system cannot be partitioned into independent subsystems of arbitrary sizes, and one has to treat each dynamical scale separately, as discussed next.

III. THE MULTICANONICAL DISTRIBUTION

Here we consider a multiscale system of size L in thermal equilibrium at temperature T_0 . We assume that the system possesses a hierarchy of dynamical structures of characteristic sizes $\ell_i = L/b^{i-1}$, with $i = 1, 2, \dots, n$, where b is a number greater than 1. (The specific value

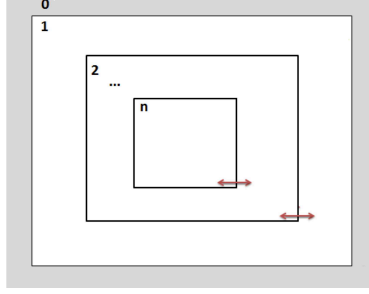


FIG. 2. A multicanonical system in thermal equilibrium at temperature T_0 .

of b is not relevant here.) We suppose furthermore that there is a wide separation of time scales within this hierarchy, with smaller structures having shorter characteristic times. Let us now consider a partition of our system into “nested” subsystems of sizes ℓ_j , as indicated in Fig. 2. We shall designate the subsystem of size ℓ_j by the label j . The “thermodynamic state” of each subsystem j will be characterized by only one parameter, namely, its inverse temperature β_j . As before, we designate the combined system consisting of all subsystems by the label 0.

Let us now focus our attention on the subsystem n of size ℓ_n . Since ℓ_n is the smallest characteristic length scale in the system, it is clear that this subsystem can be arbitrarily divided into two independent subsystems. Thus, repeating the same reasoning that led to Eq. (8), one obtains that the energy distribution law for this subsystem is given by

$$p(E|\beta_n) = \frac{g(E) \exp(-\beta_n E)}{Z_0(\beta_n)}. \quad (10)$$

Note, however, that owing to the (intermittent) energy exchange between subsystem n and its immediate surrounding, represented by subsystem $n - 1$, the parameter β_n is no longer constant but rather will fluctuate randomly. If we denote by $f(\beta_n)$ the probability density function (PDF) of β_n , then the marginal distribution $p(E)$ for subsystem n reads

$$p(E) = g(E) \int_0^\infty \frac{\exp(-\beta_n E)}{Z_0(\beta_n)} f(\beta_n) d\beta_n. \quad (11)$$

Next we wish to compute $f(\beta_n)$. To do so, we shall make use of Bayesian analysis [9]. First recall that subsystem n is embedded in a much larger subsystem $n - 1$, characterized by the parameter β_{n-1} , which is assumed to vary much slower than β_n . We are thus interested in computing $f(\beta_n|\beta_{n-1})$. From Bayes’ theorem [9] one has

$$f(\beta_n; \beta_{n-1}|E) \propto p(E|\beta_n) f(\beta_n; \beta_{n-1}), \quad (12)$$

where β_{n-1} is considered a (hyper)parameter of the distribution of β_n . In Bayesian parlance, the distribution $f(\beta_n; \beta_{n-1})$ is called the prior distribution, $p(E|\beta_n)$ is the likelihood function, and $f(\beta_n; \beta_{n-1}|E)$ is the posterior distribution. [In Eq. (12) we introduced the notation $f(\beta_n; \beta_{n-1}) \equiv f(\beta_n|\beta_{n-1})$ for convenience.]

Let us assume, as is often done in Bayesian analysis, that the prior distribution is conjugate to the likelihood $p(E|\beta_n)$, meaning that the posterior distribution follows the same parametric form as the prior distribution. If we consider the rather general case where $g(E) \propto E^{\gamma-1}$, $\gamma > 0$, so that $Z_0(\beta_n) \propto \beta_n^{-\gamma}$, it then follows from Eq. (8) that $p(E|\beta_n)$ is given by

$$p(E|\beta_n) \propto \beta_n^\gamma E^{\gamma-1} \exp(-\beta_n E). \quad (13)$$

which when viewed as the likelihood of the parameter β_n is proportional to a gamma distribution. Now, it is well known [9] that in this case the conjugate prior is also a gamma distribution, and so we have

$$f(\beta_n|\beta_{n-1}) = \frac{1}{\beta_n \Gamma(\alpha + 1)} \left(\frac{\alpha \beta_n}{\beta_{n-1}} \right)^{\alpha+1} \exp\left(-\frac{\alpha \beta_n}{\beta_{n-1}}\right), \quad (14)$$

where α is a constant. In obtaining Eq. (14) we also used the fact that $\langle \beta_n | \beta_{n-1} \rangle = \beta_{n-1}$, as it should, since subsystem $n-1$ acts as a heat reservoir for subsystem n . By scale invariance, we assume that the distribution $f(\beta_j|\beta_{j-1})$, for $j = 1, \dots, n$, has the same form as in Eq. (14).

We now have

$$f(\beta_n) = \int_0^\infty f(\beta_n|\beta_{n-1}) f(\beta_{n-1}) d\beta_{n-1}. \quad (15)$$

Using this relation recursively then yields

$$f(\beta_n) = \int_0^\infty \dots \int_0^\infty \prod_{j=1}^n f(\beta_j|\beta_{j-1}) d\beta_1 \dots d\beta_{n-1}, \quad (16)$$

with $f(\beta_j|\beta_{j-1})$ given by Eq. (14). After inserting Eq. (16) into Eq. (11), and performing a sequence of changes of variables of the type $x_j = \alpha \beta_j / \beta_{j-1}$, one can show that the resulting multidimensional integral can be expressed in terms of known higher transcendental functions:

$$p(E) = \frac{g(E)}{Z_n} {}_nF_0(\alpha + \gamma + 1, \dots, \alpha + \gamma + 1; -\beta_0 \alpha^{-n} E), \quad (17)$$

where ${}_nF_0(\alpha_1, \dots, \alpha_n; -z)$ is the generalized hypergeometric function of order $(n, 0)$ [10]. The small-scale partition function, Z_n , is given by

$$Z_n = Z_0(\beta_0) \left[\frac{\alpha^\gamma \Gamma(\alpha + 1)}{\Gamma(\alpha + \gamma + 1)} \right]^n. \quad (18)$$

One important property of the generalized hypergeometric (GHG) distribution given in (17) is that it exhibits power-law tails of the form: $p(E) \propto E^{-(\alpha+2)}$, for $E \rightarrow \infty$. This follows immediately from the asymptotic expansion of the function ${}_nF_0$ [11]: ${}_nF_0(\alpha_1, \dots, \alpha_n; -x) = \sum_{i=1}^n C_i x^{-\alpha_i} (1 + O(1/x))$, as $x \rightarrow \infty$. It is also worth pointing out that the first two members of the family ${}_nF_0$ yield elementary functions, namely, ${}_0F_0(x) = \exp(x)$ and ${}_1F_0(1/(q-1), x) = \exp_q(x/(q-1))$, where $\exp_q(x)$ is the q -exponential: $\exp_q(x) = [1 + (1-q)x]^{1/(1-q)}$. The GHG distribution with $n = 0$ thus recovers the Boltzmann-Gibbs distribution, whereas for $n = 1$ it gives the q -exponential or Tsallis distribution [7]. One then sees from the preceding discussion that if a system with only one time scale is in thermal equilibrium then the Boltzmann-Gibbs distribution follows, whereas if it has two distinct time scales the Tsallis distribution should be applicable. For complex systems with more than two characteristic time scales, such as turbulent flows, GHG distributions of higher order are thus required [5, 6].

IV. CONCLUSIONS

We have presented an alternative derivation, based on Bayesian analysis, of the multicanonical distribution, which describes the statistical equilibrium of complex systems possessing a hierarchy of time and length scales. We have shown that the multicanonical distribution can be written explicitly in terms of generalized hypergeometric functions, which exhibit a power-law asymptotic behavior. This thus shows that the emergence of power law distributions—an ubiquitous feature in nature—is intimately connected with the existence of multiple time and length scales in the system.

ACKNOWLEDGMENTS

This work was supported in part by the Brazilian agencies CNPq and FACEPE.

-
- [1] A. Clauset, C.R. Shalizi, and M.E.J. Newman, SIAM Review **51**, 661 (2009).
 - [2] E. Alvarez-Lacalle, B. Dorow, J.-P. Eckmann, and E. Moses, Proc. Natl. Acad. Sci. USA **103**, 7956 (2006).

- [3] R. A. Treumann and C. H. Jaroschek, Phys. Rev. Lett. **100**, 155005 (2008).
- [4] An extensive bibliography on power-law distributions can be found in <http://www.nslj-genetics.org/wli/zipf>.
- [5] D. S. P. Salazar, and G. L. Vasconcelos, Phys. Rev. E **82**, 047301(2010).
- [6] D. S. P. Salazar, and G. L. Vasconcelos, “Multicanonical distribution: Statistical equilibrium of multiscale systems,” submitted to Phys. Rev. Lett. (2012).
- [7] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
- [8] G. H. Wannier, *Statistical Physics* (Dover, New York 1966)
- [9] A. Gelman, A., J. B. Carlin, H. S. Stern, and D. Rubin, *Bayesian Data Analysis*, 2nd ed. (Chapman and Hall, London, 2003).
- [10] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.
- [11] See, e.g., <http://functions.wolfram.com/07.31.06.0041.01>